

Multifractal modeling of short-term interest rates

Martin Rypdal and Ola Løvsetten

Department of Mathematics and Statistics, University of Tromsø, Norway

November 2011

We propose a multifractal model for short-term interest rates. The model is a version of the Markov-Switching Multifractal (MSM), which incorporates the well-known level effect observed in interest rates. Unlike previously suggested models, the level-MSM model captures the power-law scaling of the structure functions and the slowly decaying dependency in the absolute value of returns. We apply the model to the Norwegian Interbank Offered Rate with three months maturity (NIBORM3) and the U.S. Treasury Bill with three months maturity (TBM3). The performance of the model is compared to level-GARCH models, level-EGARCH models and jump-diffusions. For the TBM3 data the multifractal out-performs all the alternatives considered.

1. Introduction

Interest rates play an important role for financial institutions, for instance in risk management. Interest-rate risk is often assessed via simple stress tests, where one considers parallel shifts in the yield curve (typically 100 or 200 basis points) and reports the changes in the value of the portfolio. This approach can be improved by stochastic modeling of future movements in the interest-rate yield curve. Vasicek [1977] and Cox et al. [1985] argue that the yield curve is given by the spot rate alone. Short-term interest rates, such as NIBORM3 and TBM3, are frequently used as proxies for the spot rate, and hence accurate modeling of these time series is potentially very important. To our knowledge the present paper is the first published study considering the Norwegian rate, while the TBM3 has been studied by e.g. Andersen and Lund [1997], Chapman and Pearson [2001], Durham [2003], Johannes [2004] and Bali and Wu [2006].

Traditionally, short-term rates have been modeled by Itô stochastic differential equations on the form

$$dR(s) = f(R(s))ds + cR(s)^\gamma dB(s), \quad (1)$$

where s is the time variable, $R(s)$ denotes the risk-free rate and $B(s)$ is a Brownian motion. If we discretize this equation, by letting $s = t\Delta s$ and $r_t = R(s)$, then we obtain a stochastic difference equation on the form $r_t = r_{t-1} + f(r_{t-1})\Delta t + c\Delta s^{1/2}r_{t-1}^\gamma w_t$, where w_t are independent and Gaussian distributed random variables with unit variance. It is convenient to write this equation on the form

$$r_t = \mu_t + \sigma_t, \quad (2)$$

with $\mu_t = r_{t-1} + \Delta s f(r_{t-1})$ and $\sigma_t = r_{t-1}^\gamma x_t$. Here $x_t = \sigma w_t$ with $\sigma = c\Delta s^{1/2}$. Throughout the paper we will consider linear drift terms on the form $f(r) = A_0 + A_1 r$. If $A_1 < 0$, then (1) has a stationary solution, and the discretized equation has a stationary solution for sufficiently small $\Delta s > 0$. For a fixed $\Delta s > 0$ we write $\mu_t = r_{t-1} + \alpha_0 + \alpha_1 r_{t-1}$. In this case we have stationarity for $-2 < \alpha_1 < 0$.

The number $\gamma \geq 0$ is called the Constant Elasticity Variance (CEV) parameter. For $\gamma \neq 0$ the models feature the so-called level effect. It is generally accepted that this effect is present in interest-rate data, see for instance [Longstaff et al., 1992]. The level effect introduces volatility persistence, i.e. strong dependence

between the absolute values of the increments $\Delta r_t = r_t - r_{t-1}$. However, if $B(s)$ is a Brownian motion, then the variables

$$\frac{\Delta r_t - (\alpha_0 + \alpha_1 r_{t-1})}{r_{t-1}^\gamma} \quad (3)$$

are i.i.d. and Gaussian, and hence the volatility persistence will vanish under a simple transformation of the increments. Such processes are called pure-level models and have been studied by Cox et al. [1985] and Longstaff et al. [1992]. Using time-series data for short-term interest rates we can optimize the likelihood to determine the parameters α_0 , α_1 and γ . These results are shown in table 6. We can then transform the data according to (3). The resulting time series (which are plotted in figures 1(c) and 2(c)) are realizations of the so-called normalized increment process. Just from inspection of figures 1(c) and 2(c) we can observe that the resulting time series exhibit strong volatility persistence. This is confirmed in figures 1(a) and 1(b), where we have plotted the autocorrelation functions for the absolute values of the normalized increments. As a consequence of these observations we conclude that pure-level models are insufficient, and we will therefore replace the process w_t with dependent variables x_t .

In our empirical investigations we find that the dependency structure in the variables x_t resembles the stylized facts of logarithmic returns of asset prices, namely that the variables themselves are uncorrelated (or weakly correlated) whereas their absolute values have slowly decaying autocorrelation functions. To describe this dependency several authors (e.g. Brenner et al. [1996] and Koedijk et al. [1997]) have suggested using a generalized autoregressive conditional heteroskedasticity (GARCH) model for the process x_t . The corresponding process r_t is then called a level-GARCH model. These models can be further improved by letting x_t be an Exponential GARCH (EGARCH) model or a Markov Switching GARCH (MSGARCH) model. The level-EGARCH model was introduced by Andersen and Lund [1997] and has been reported to perform better than the standard GARCH model on the TBM3 data.

In this work we propose a multifractal model, specifically a level-MSM model, as an alternative to level-EGARCH models and level-MSGARCH models for short-term interest rates. This model is a slight modification of the standard MSM model constructed by Calvet and Fisher (2004). The motivation for introducing multifractal models for short-term interest rates is similar to the motivation for multifractal modeling of asset prices, namely that these models capture the dependency structure and scaling properties of the process x_t . Secondly, encouraging results (in- and out-of-sample) have been obtained in a preliminary study of the NIBORM3 data [Løvsetten, Autumn 2010].

We follow Calvet and Fisher (2004) and use an adjusted version of the Vuong test [Vuong, 1989] for model selection, and the main result of this work is that the level-MSM model performs well compared against level-EGARCH models, level-MSGARCH models and jump-diffusions. The paper is organized as follows: In section 2 we define the level-MSM model and consider some basic properties. Technical details about multifractal processes are presented in appendix 1. A brief description of the alternative models are presented in section 3, and in section 4 the results of the Vuong test are presented. In section 5 we make some concluding remarks.

Remark 1.1. The data analyzed in this paper is freely available online. Both data sets are given with daily resolution. The TBM3 data is taken from the period 1951-01-04 to 2010-09-22, and consists of 14172 data points. The NIBORM3 data is taken from the period 1986-01-02 to 2010-09-24, and consists of 6231 data points. The TBM3 data contains several days with zero value, which causes certain technical problems in the stochastic modeling. The problem is resolved by a simple variable shift $r \mapsto r + b$. In order to be consistent with our set of models, we estimate the constant $b > 0$ by estimating the conditional standard deviations $\text{sdv}[\Delta r_t | r_{t-1} = r]$ (for various r) and fitting a function $cr^\gamma - b$ to this data set. With this approach we find $b = 0.03$.

2. Multifractal models

The application of multifractal processes to finance was introduced by Mandelbrot et al. [1997]. A process $X(t)$ is multifractal if the structure functions $S_q(t) = \mathbb{E}[|X(t)|^q]$ are power laws in t , and one define a

scaling function $\zeta(q)$ through the relation $S_q(t) \sim t^{\zeta(q)}$ as $t \rightarrow 0$, i.e.

$$\zeta(q) = \lim_{t \rightarrow 0} \frac{\log S_q(t)}{\log t}.$$

Using Hölder's inequality it is easy to show that the scaling function $\zeta(q)$ is concave, and we also note that if the process $X(t)$ is h -self similar, i.e. $X(at) \stackrel{d}{=} a^h X(t)$, then $\zeta(q) = hq$. We are therefore interested in the situations where $\zeta(q)$ is strictly concave. In this case we see that

$$\frac{\mathbb{E}[X(t)^4]}{\mathbb{E}[X(t)^2]^2} \sim t^{\zeta(4)-2\zeta(2)} \rightarrow \infty \quad \text{as } t \rightarrow 0.$$

If we assume that $X(t)$ also has stationary increments, then this implies that the Δt -lagged increments $X(t + \Delta t) - X(t)$ are more leptokurtic for small Δt than for larger Δt . In particular, $X(t)$ can not be a Gaussian process.

In financial time series one will most often want the process $X(t)$ to have uncorrelated and stationary increments. In order for this to be satisfied we must impose the condition $\zeta(2) = 1$, because otherwise the variables $\Delta X(t) = X(t + 1) - X(1)$ have slowly decaying autocorrelation. In fact, for $\zeta(2) \neq 1$

$$\text{Cov}(\Delta X(0), \Delta X(t)) \sim t^{2H_d-2},$$

where $H_d := \zeta(2)/2$. The advantage of a multifractal process (with strictly concave scaling function) is that it has strongly dependent increments, even for $H_d = 1/2$. For instance, in the MSM model¹ we have $\zeta(2) = 1$ and

$$\text{Cov}(|\Delta X(0)|^q, |\Delta X(t)|^q) \sim t^{-(\zeta(2q)-2\zeta(q))}. \quad (4)$$

See proposition 1 in [Calvet and Fisher, 2004]. This inherent long-range volatility persistence serves as our motivation for modeling short-term interest rates using multifractals.

2.1. Mandelbrot's MMAR processes

The multifractal models introduced by Mandelbrot et al. [1997] are stochastic processes on the form

$$X(t) = \sigma T^{1/2} B(\Theta(t)), \quad (5)$$

defined for $0 \leq t \leq T$, where $\Theta(t) = \mu([0, t])$ is a random probability measure on $[0, T]$ and $B(t)$ is a Brownian motion with $\mathbb{E}[B(1)^2] = 1$. The process $\Theta(t)$ is itself a multifractal process with scaling function $\zeta_\Theta(q)$, and if μ and $B(\cdot)$ are independent, then $\zeta_X(q) = \zeta_\Theta(q/2)$. By construction, the measures we will mention in this paper satisfy $\mathbb{E}[\mu([0, t])] \propto t$, i.e. $\zeta_\Theta(1) = 1$. It follows that $\zeta_X(2) = 1$. We note that processes with correlated increments are useful in other applications (see e.g. [Rypdal and Rypdal, 2010]), and the Brownian motion $B(t)$ can in these situations be replaced by a fractional Brownian motion with $H \neq 1/2$.

There are several choices for the measure μ , and for completeness we give three examples in the appendix. We also show how a particular construction, the Poisson multifractal, leads to a discrete-time MSM model defined in the next section.

2.2. The MSM model

The Markov Switching Multifractal model of order $K > 1$ is given by

$$x_t = \sigma g(\mathbf{M}_t)^{1/2} \varepsilon_t, \quad (6)$$

¹It is pointed out by Lux [2006] that the MSM model is only a finite-level approximation, and so equation (4) is only valid on time scales up to b^K . Here $b > 0$ and $K \in \mathbb{N}$ are parameters in the MSM model. See section 2.2.

where $\mathbf{M}_t = (M_{1,t}, \dots, M_{K,t})$, and the function $g : \mathbb{R}^K \rightarrow \mathbb{R}$ is the product of the vector components. The innovations ε_t are Gaussian distributed with unit variance, and in this setting also assumed to be independent. The process \mathbf{M}_t is a Markov chain defined by the following updating scheme:

$$M_{k,t} = \begin{cases} \text{drawn from } M & \text{with probability } \lambda_k, \\ M_{k,t-1} & \text{with probability } 1 - \lambda_k. \end{cases} \quad (7)$$

Components are updated independently of all previous updates, and the frequencies λ_k are related to each other through (A.4).

For simplicity we choose a two-point distribution M :

$$\mathbb{P}(M = m_0) = \mathbb{P}(M = m_1) = \frac{1}{2}, \text{ with } m_0 \in (1, 2) \text{ and } m_1 = 2 - m_0. \quad (8)$$

This version of the MSM model is known as the binomial multifractal [Calvet and Fisher, 2004]. To highlight the number of multipliers K we will use the notation $\text{MSM}(K)$. Note that, unlike GARCH(p, q), the number of parameters remains unchanged with increasing order.

Let $\{\mathbf{m}^i | i = 1, \dots, d := 2^K\}$ denote the sample space of \mathbf{M}_t . We also define vectors $\omega(x) : \mathbb{R} \rightarrow \mathbb{R}^d$ and $\mathbf{p}_t : \mathbb{N} \rightarrow \mathbb{R}^d$ with components $\omega_k(x) = n(x|\sigma^2 g(\mathbf{m}^k))$ and $p_{k,t} = \mathbb{P}(\mathbf{M}_t = \mathbf{m}^k | \mathbf{x}_t)$ respectively. Here $\mathbf{x}_t = (x_2, \dots, x_t)$, and $n(\cdot|\sigma)$ is the density of the normal distribution with zero mean and variance σ^2 . With this notation the density of $x_t | \mathbf{x}_{t-1}$ is given by

$$p(x_t | \mathbf{x}_{t-1}) = \mathbf{p}_{t-1}^T A \omega(x_t), \quad (9)$$

where A is the transition matrix of the Markov chain, that is $A_{j,k} = \mathbb{P}(\mathbf{M}_t = \mathbf{m}^k | \mathbf{M}_{t-1} = \mathbf{m}^j)$. This is seen by conditioning on the underlying Markov-chain at time $t-1$ and t , and using Bayes' rule. It is easily seen, once again using Bayes' rule, that the vectors \mathbf{p}_t follow the recursion

$$\mathbf{p}_t^T = \frac{1}{p(x_t | \mathbf{x}_{t-1})} \omega^T(x_t) * (\mathbf{p}_{t-1}^T A), \quad (10)$$

where $*$ denotes the Hadamard product. We start the recursion with the limiting probabilities of \mathbf{M}_t .

2.3. The level-MSM model for interest rates

We propose the following level-MSM model for short-term interest rates:

$$r_t - r_{t-1} = \alpha_0 + \alpha_1 r_{t-1} + r_{t-1}^\gamma x_t, \quad (11)$$

where x_t is the binomial MSM model defined above. The likelihood

$$L = \sum_{t=2}^n \log f(r_t | \mathbf{r}_{t-1}),$$

for data r_2, \dots, r_n , with r_1 taken as a pre-sample value, now follows from equations (9) and (10) together with the relation

$$f(r_t | \mathbf{r}_{t-1}) = \frac{1}{r_{t-1}^\gamma} p(x_t | \mathbf{x}_{t-1}),$$

where f is the density of $r_t | \mathbf{r}_{t-1}$.

We have fitted the level-MSM models to the two time series under consideration for $K = 2, \dots, 9$. For the case $K = 9$, the maximum likelihood (ML) estimates for the parameters α_0 and α_1 are reported in table 6. In this table the estimates of α_0 and α_1 for the alternative models are also included. As expected the estimates for the parameter α_1 are negative, but we observe that none of these ML estimates are significantly different from zero for the TBM3 data. The estimates for α_1 are also very small for the NIBORM3 data, and we will therefore consider the model defined by (11) with $\alpha_1 = 0$. The corresponding ML estimates for the other parameters are presented in table 6. From these results we also see that the likelihoods increase monotonically with K . Using a Vuong test we compare the level-MSM of order $K = 9$ against level-MSM models of lower order. When comparing two models with this method, the null-hypothesis is that

both models are equally far from the data-generating process measured by the Kullback-Leibler distance. Hence small p -values indicate that the level-MSM model of order $K = 9$ is significantly closer to the true generating process than the level-MSM models of order $K = 7$ or lower.

The same model-selection-test is used to compare the level-MSM model with various alternative models, and more details on the test are presented in section 4.

3. Alternative models

In this section we briefly discuss some processes which we use for benchmarking the level-MSM model. As for the level-MSM model we have first considered these models with drift terms on the form $\mu_t = r_{t-1} + \alpha_0 + \alpha_1 r_{t-1}$. The ML estimates for the parameters α_0 and α_1 are reported in table 6. Again we see that the contribution from the parameter α_1 is negligible. Consequently we will consider models with drift terms on the form $\mu_t = r_{t-1} + \alpha_0$.

3.1. The level-GARCH model

In this paper we prefer the standard level-GARCH model proposed by Koedijk et al. [1997]. As is common for this model, we include student- t distributed innovations. The model then reads

$$r_t = \alpha_0 + r_{t-1} + \sigma_t, \quad \sigma_t = h_t^{1/2} r_{t-1}^\gamma \varepsilon_t, \quad h_t = a_0 + a_1 x_{t-1}^2 + b h_{t-1}, \quad (12)$$

where the innovations ε_t are i.i.d. $t_\nu(0, 1)^2$. For $\gamma = 0$ we have the pure-GARCH(1,1) model, and for $a_1 = b = 0$ we have the standard level model.

The estimated parameters for the NIBORM3 and TBM3 data are presented in table 6. We observe that simulated paths with these exponents have far too wild fluctuations compared to the real data, indicating that this model fails to accurately describe the interest-rate fluctuations. In addition, we know that the GARCH models exhibit exponentially decaying autocorrelation functions for the absolute values of the increments. This means that the long-range volatility persistence observed in the short-term interest rates is not inherent in these models.

3.2. The level-EGARCH model

Instead of using (12), Andersen and Lund [1997] propose using the EGARCH model. They find that this model gives an adequate fit to the TBM3 data, and a better fit compared to the standard GARCH model. In the EGARCH model the logarithm of the conditional variance replaces the conditional variance. The variance-recursion is then

$$\log h_t = a_0 + a_1 \varepsilon_{t-1} + a_2 |\varepsilon_{t-1}| + b \log h_{t-1}. \quad (13)$$

The extra parameter a_1 controls potentially different responses to positive and negative returns. The use of the logarithm guarantees positive values of the volatility for all parameter values. In addition, this model provides the extra flexibility by letting the conditional distribution be non-symmetric. As we will become apparent from the results presented in the next section, the EGARCH model gives better results than the GARCH model for both of the time series considered in this paper. This confirms the results of Andersen and Lund [1997].

The ML estimates for the parameters in the EGARCH model are presented in table 6.

3.3. Jump-diffusions

The final class of alternative models considered are the jump-diffusions. Both Johannes [2004] and Das [2002] propose these processes in order to describe the large spikes observed in interest-rate data. As a discretized jump-diffusion model we use the following specification:

$$r_t = \alpha_0 + r_{t-1} + r_{t-1}^\gamma (x_t + J_t z_t),$$

² $t_\nu(0, 1)$ is a centralized, unit-variance student- t distribution with ν degrees of freedom.

where J_t is the jump-indicator assumed to follow a Bernoulli-distribution. The probability of a jump taking place at time t is then given by $(1 + \exp(-c - dr_{t-1}))^{-1}$. By letting $x_t = h_t^{1/2} \varepsilon_t$ follow a GARCH-process, conditional heteroschedacity in levels is also accomodated [Das, 2002; Hong et al., 2004]. The innovations are distributed as $z_t \stackrel{d}{\sim} \mathcal{N}(0, \tau^2)$ and $\varepsilon_t \stackrel{d}{\sim} \mathcal{N}(0, 1)$. The variance recursion in the GARCH process is given by

$$h_t = a_0 + a_1 \left(\frac{\Delta r_{t-1} - \alpha_0}{r_{t-2}^\gamma} \right)^2 + b h_{t-1},$$

where $\Delta r_t = r_t - r_{t-1}$.

The ML estimates for the parameters in the jump-diffusion model are presented in table 6.

4. In-sample comparison

To test the binomial MSM model against the alternative models we emply a version of the Vuong test which is adjusted for heteroschedacity and autocorrelation (HAC) [Calvet and Fisher, 2004]. For each of the alternative models the null-hypothesis is that this model and the level-MSM model of order $K = 9$ are equidistant from the true data-generating process, measured in the Kullback-Leibler distance [Kullback and Leibler, 1951]. In the classical Vuong test it is assumed that the data-generating process is i.i.d., and then the log-ratio of the likelihoods for the two models will converge to a normal distribution with zero mean, and with a variance which is consistently estimated by the sample variance for the log-ratio of the likelihoods. Using the corresponding normal distribution one can then easily calculate a p -value under the null-hypothesis. In the HAC-adjusted version of this test, the data-generating process may exhibit dependence, but the variables r_t should be identically distributed. This is satisfied if we assume that the data-generating process r_t is stationary.

The results of this test are presented in table 6. We observe that for the TBM3 data, the level-MSM model of order $K = 9$ is significantly closer to the data-generating process than any of the alternative models. For the NIBORM3 data the EGARCH model performs best, whereas the level-MSM model performs better than the standard level-GARCH model and jump-diffusions.

5. Concluding remarks

In this paper we have introduced a multifractal model for short-term interest rates. The model combines the well-established level effect described in [Longstaff et al., 1992] with the discretized multifractal model of [Calvet and Fisher, 2001]. In a comparison with level-GARCH, level-EGARCH and jump-diffusions, we find that this model well describes the fluctuations of the TBM3 and NIBORM3 time series. The main result of this work is that the level-MSM outperforms all alternatives for the TBM3 data. This motivates further research on multifractal modeling of short-term interest rates, in particular an out-of-sample analysis of the level-MSM model.

It is also interesting to note in the level-GARCH model the parameter estimates (for both the TBM3 data and the NIBORM3 data) fall outside the covariance-stationarity region. As a result the level-GARCH model has a wild volatility pattern, which does not seem to be an accurate description of the interest-rate data. This confirms the results of e.g. Andersen and Lund [1997].

Appendix A Examples of multifractal measures

In this appendix we give examples of random measures that can be used to define multifractal processes via (5). For the Poisson multifractal we show how a discretization leads to the model in (6).

Example 1: In the simplest case μ is a randomized dyadic Bernoulli measure with probabilities $p_1 = p$ and $p_2 = 1 - p$. This measure is constructed through an iterative procedure, where we in the first step divide the interval $[0, T]$ in two pieces Δ_0 and Δ_1 of equal length. One of the intervals is chosen at random

and given a mass $p_1 \mu([0, T])$, while the other interval is given the mass $p_2 \mu([0, T])$. This procedure is then repeated recursively, i.e. the interval Δ_0 is divided into the equally sized intervals Δ_{00} and Δ_{01} . One of these intervals is given the mass $p_1^2 \mu([0, T])$ and the other is given the mass $p_1 p_2 \mu([0, T])$. Formally the measure can be defined by letting f be a random bijection³ $\{0, 1\} \rightarrow \{p_1, p_2\}$ and f_{i_1, \dots, i_k} be independent copies of f . The measure is constructed by assigning the mass $\mu(\Delta_{i_1, i_2, \dots, i_k}) = f(i_1) f_{i_1}(i_2) \cdots f_{i_1, i_2, \dots, i_{k-1}}(i_k)$ to the dyadic intervals

$$\Delta_{i_1, \dots, i_k} = [T \cdot (0.i_1 \cdots i_k)_2, T \cdot (0.i_1 \cdots i_k)_2 + T \cdot 2^{-k}], \quad i_n \in \{0, 1\}.$$

A simple combinatory argument shows that for $t = T \cdot 2^{-k}$, the random variable $\Theta(t)$ has density

$$p_{\Theta, t}(x) = \frac{1}{2^k} \sum_{n=0}^k \binom{k}{n} \delta(x - p^n (1-p)^{k-n}).$$

When the processes $B(t)$ and $\Theta(t)$ are independent, we obtain the density of $X(t)$:

$$p_{X, t}(x) = \int p_{\sigma T^{1/2} B, s}(x) p_{\Theta, t}(s) ds = \frac{1}{2^k} \sum_{n=0}^k \binom{k}{n} \frac{\exp\left(-\frac{x^2}{2\sigma^2 (T p^n (1-p)^{k-n})}\right)}{\sqrt{2\pi\sigma^2 (T p^n (1-p)^{k-n})}}.$$

From this density we can easily calculate the structure functions $S_q(t) = \mathbb{E}[|X(t)|^q]$ and see that $S_q(t) \sim t^{\zeta(q)}$ as $t \rightarrow 0$, where $\zeta(q) = 1 - T(q/2)$ and $T(q) = \log_2(p^q + (1-p)^q)$.

Remark 5.1. The function $T(q)$ is sometimes called the scaling function of the random measure. The dyadic measure defined by $\mu(\Delta_{i_1, \dots, i_k}) = p_{i_1} \cdots p_{i_k}$ has a multifractal spectrum $f(\alpha)$ given by the Legendre transform of $T(q)$: $f(\alpha) = \inf_{q \in \mathbb{R}} \{\alpha q - T(q)\}$, and the Hentchel-Proccaccia dimension spectrum is $D_q = (q-1)^{-1} T(q)$. See e.g. [Pesin, 1997] for an account of the relation between scaling functions and spectra of fractal dimensions. The function $\alpha \mapsto f(\alpha/2)$ is often called the singularity spectrum of the process $X(t)$.

Example 2: A different class of multifractal measures are the b -adic random multiplicative cascades (see e.g. [Mandelbrot et al., 1997] for a more detailed account). For an integer $b \geq 2$ and $i_n \in \{0, 1, \dots, b-1\}$ we define the b -adic subintervals of $[0, T]$ as

$$\Delta_{i_1, \dots, i_k} = [T \cdot (0.i_1 \dots i_k)_b, T \cdot (0.i_1 \dots i_k)_b + T \cdot b^{-k}].$$

For a positive random variable M , with $\mathbb{E}[M] = b^{-1}$, let

$$\mu(\Delta_{i_1, \dots, i_k}) = M_{i_1} M_{i_1 i_2} \cdots M_{i_1 i_2 \dots i_k} \Omega_{i_1, \dots, i_k},$$

where M_{i_1, \dots, i_n} are independent copies of M and

$$\Omega_{i_1, \dots, i_k} = \lim_{n \rightarrow \infty} \sum_{j_1, \dots, j_n} M_{i_1} \cdots M_{i_1 i_2 \dots i_k} M_{i_1 i_2 \dots i_k j_1} \cdots M_{i_1 i_2 \dots i_k j_k \dots j_n}.$$

Again a stochastic process is constructed according to equation (5) with $\Theta(t) = \mu([0, t])$. This process has a scaling function $\zeta(q) = 1 - T(q)$, where $T(q) = \log_b \mathbb{E}[M^q]$. Popular choices for the multiplier M are log-normal distributions, which give quadratic scaling functions, or other log-infinitely divisible distributions. Almost every realization of the measure μ has the multifractal spectrum $f(\alpha) = \inf_{q \in \mathbb{R}} \{\alpha q - T(q)\}$.

Example 3: The Poisson multifractal measure on $[0, T]$ generalizes the b -adic multiplicative cascade by introducing randomness in the construction of the intervals $\Delta_{i_1, i_2, \dots, i_k}$. In the original multifractal models, each interval on level k are divided into b pieces of equal length at level $k+1$. As a result the interval length decreases as b^{-k} with the level k . In the Poisson multifractal measure, the splitting of an interval at level k is performed by drawing the lengths of the new pieces randomly from an exponential distribution with rate l_{k+1} . This means that the mean and median length of an interval at level $k+1$ are $1/l_{k+1}$ and

³With probability $1/2$ for each of the two outcomes $(0, 1) \mapsto (p_1, p_2)$ and $(0, 1) \mapsto (p_2, p_1)$

$\log(2)/l_{k+1}$ respectively, so to maintain exponential decay of interval lengths (as a function of level), one chooses $l_k = b^{k-1}l_1$. Note that b no longer is restricted to the integers.

Formally, the Poisson multifractal measure on the interval $[0, T]$ is defined via a sequence of measures μ_k specified on randomly generated intervals

$$\Delta_{i_1, i_2, \dots, i_k} = [t_{i_1, i_2, \dots, i_k}, t_{i_1, i_2, \dots, (i_k+1)}],$$

where the numbers t_{i_1, \dots, i_k} are defined by the following recursive construction: Let $\{\tau_{i_1, \dots, i_k}(j)\}_{j \in \mathbb{N}}$ denote independent and exponentially distributed random variables with rates $l_k = b^{k-1}l_1$. Define

$$N_{i_1, \dots, i_{k-1}} = \max \left\{ m : \sum_{j=1}^m \tau_{i_1, \dots, i_{k-1}}(j) < \text{diam}(\Delta_{i_1, i_2, \dots, i_{k-1}}) \right\} \quad (\text{A.1})$$

and

$$t_{i_1, \dots, i_k} = \begin{cases} t_{i_1, \dots, i_{k-1}} & \text{if } i_k = 0 \\ t_{i_1, \dots, i_{k-1}} + \sum_{j=1}^{i_k} \tau_{i_1, \dots, i_{k-1}}(j) & \text{if } 1 \leq i_k \leq N_{i_1, \dots, i_{k-1}} \\ t_{i_1, \dots, (i_{k-1}+1)} & \text{if } i_k = N_{i_1, \dots, i_{k-1}} + 1 \end{cases} \quad (\text{A.2})$$

This means that the interval Δ_{i_1, \dots, i_k} is divided into N_{i_1, \dots, i_k} subintervals by the cuts made by a Poisson process with rate l_{k+1} . We start with an interval $\Delta = [0, T]$. A sequence μ_k of measures can now be defined via the formula

$$\mu_k(\Delta_{i_1, i_2, \dots, i_k}) = T^{-1} \text{diam}(\Delta_{i_1, i_2, \dots, i_k}) M_{i_1} M_{i_1, i_2} \cdots M_{i_1, \dots, i_k},$$

where M_{i_1, \dots, i_k} are independent copies of a positive random variable M satisfying $\mathbb{E}[M] = 1$. If $\mathbb{E}[M^2] < b$, then μ_k converges weakly to a Borel measure μ on $[0, T]$. With this choice of random measure the model given by equation (5) has scaling function

$$\zeta(q) = 1 - T(q/2), \quad \text{where } T(q) = 1 - q + \log_b \mathbb{E}[M^q]. \quad (\text{A.3})$$

Calvet and Fisher constructed their MSM model by discretizing the time interval $[0, T]$ and assigning discrete geometric (rather than an exponential) distributions on the waiting times τ . More precisely one will fix an integer $K > 1$ determining the number of levels that are to be included in the discrete model, and consider the integer values $\{0, 1, \dots, \tilde{T}\}$, where $\tilde{T} = m^K$ for some positive integer m . As for the construction of the Poisson multifractal one makes random partitions $\Delta_{i_1, i_2, \dots, i_k} = [t_{i_1, i_2, \dots, i_k}, t_{i_1, i_2, \dots, (i_k+1)}]$ of the interval $[0, \tilde{T}]$. These are constructed using the recursive procedure described by equations (A.1) and (A.2), where the variables τ_{i_1, \dots, i_k} are replaced by discrete random variables $\tilde{\tau}_{i_1, \dots, i_k}$ with geometric distributions $\mathbb{P}(\tau_{i_1, \dots, i_k} = \tau') = \lambda_k (1 - \lambda_k)^{\tau'-1}$, with $1 - \lambda_k = e^{-lb^{k-1}\tilde{T}/T}$. This choice of parameters λ_k implies that the median of $\tilde{\tau}_{i_1, \dots, i_k}$ is proportional to $l^{-1}b^{-n}$. Note that the parameters λ_k are given by λ_1 through the formula

$$\lambda_k = 1 - (1 - \lambda_1)^{b^{n-1}}. \quad (\text{A.4})$$

A measure $\tilde{\mu}$ is defined by specifying that if t is an integer and $[t-1, t] \subseteq \Delta_{i_1, \dots, i_K}$, then

$$\tilde{\mu}([t-1, t]) = m^{-K} M_{i_1} M_{i_1, i_2} \cdots M_{i_1, \dots, i_K},$$

where M_{i_1, \dots, i_k} are independent copies of a positive random variable M with $\mathbb{E}[M] = 1$. Again we define $\tilde{\Theta}(t) = \tilde{\mu}([0, t])$, and by taking the composition with a fractional Brownian motion we get a discrete-time stochastic process $\tilde{X}(t) = C B(\tilde{\Theta}(t))$. We denote $x_t = \tilde{X}(t) - \tilde{X}(t-1)$, and observe that

$$x_t \stackrel{d}{=} C \tilde{\mu}([t, t-1])^{1/2} (B(t) - B(t-1)) = \sigma (M_{i_1} M_{i_1, i_2} \cdots M_{i_1, \dots, i_K})^{1/2} \varepsilon_t,$$

where $\sigma = C m^{-K}$ and $\varepsilon_t = B(t) - B(t-1)$ is a discrete version of a white Gaussian noise. We can simplify notations by denoting $M_{k,t} = M_{i_1, \dots, i_k}$ for $t \in \Delta_{i_1, \dots, i_k}$.

We finally remark that there exists a large class of multifractal random measures, known as log-infinitely divisible cascades [Bacry et al., 2001; Bacry and Muzy, 2003; Bacry et al., 2008], which have multifractal scaling and stationary increments. However, there only exists approximate maximum likelihood

methods for these processes [Løvsletten and Rypdal, 2011], and hence the Young-testing performed in this paper is not available.

Acknowledgment. This project was partly funded by *Sparebank 1 Nord-Norge* and the Norwegian Research Council (project number 208125).

6. Figures and tables

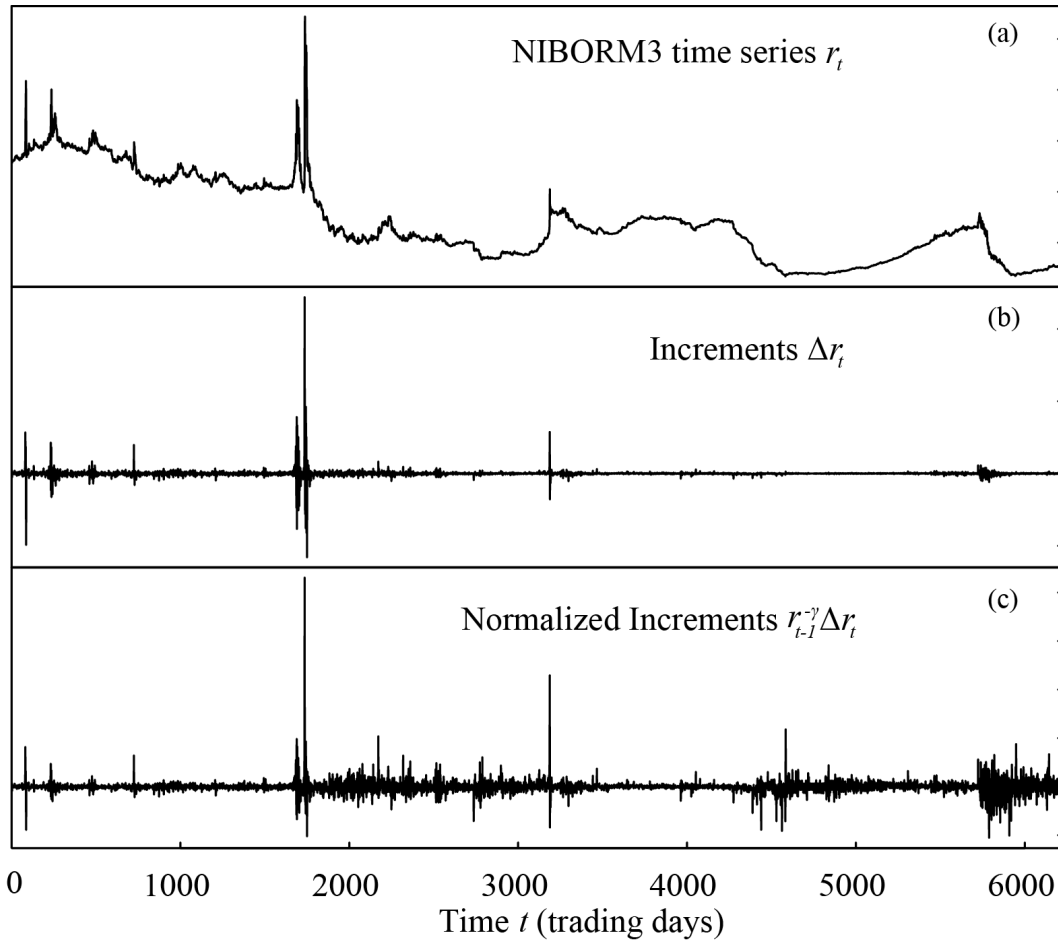


Fig. 1. (a): The NIBORM3 data r_t for the time period 1986-01-02 to 2010-09-24. (b): The one-day increments $\Delta r_t = r_t - r_{t-1}$. (c): The normalized one-day increments $\Delta r_t / r_{t-1}^\gamma$. The value of the CEV parameter is $\gamma = 1.61$.

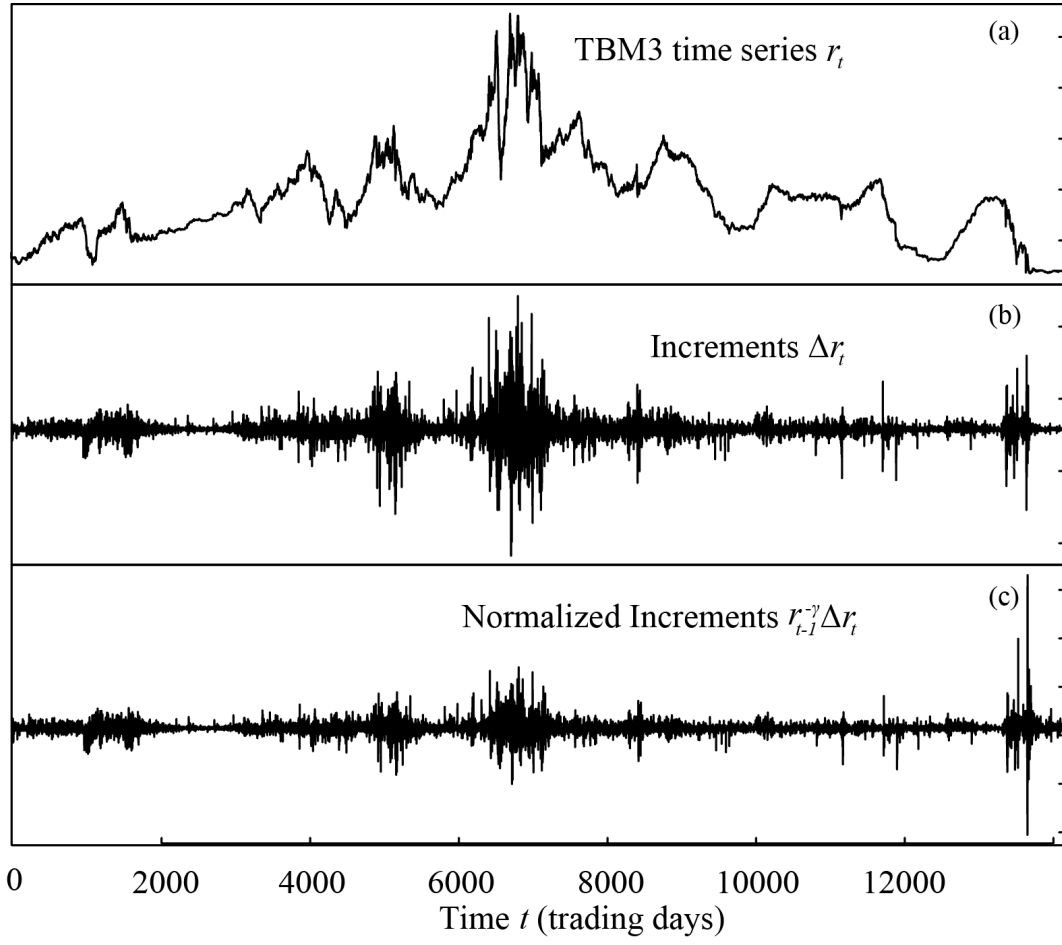


Fig. 2. (a): The TBM3 data r_t for the time period 1954-01-04 to 2010-09-22. (b): The one-day increments $\Delta r_t = r_t - r_{t-1}$. (c): The normalized one-day increments $\Delta r_t / r_{t-1}^\gamma$. The value of the CEV parameter is $\gamma = 0.38$.

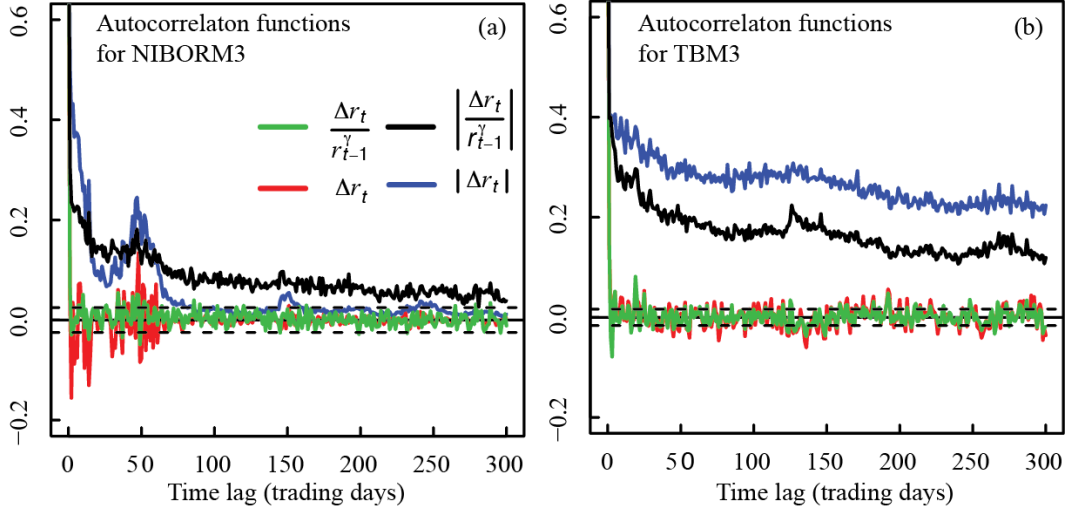


Fig. 3. (a): Autocorrelation functions for one-day increments in the NIBORM3 data. The variables considered are the standard increments $\Delta r_t = r_t - r_{t-1}$, the normalized increments $\Delta r_t / r_{t-1}^\gamma$ and the absolute values of these two time series. The value of the CEV parameter is $\gamma = 1.61$. (b): Same as (a), but now for the TBM3 data. The value of the CEV parameter is $\gamma = 0.38$. In both figure the dotted lines represent a 0.95 confidence interval for the autocorrelation function assuming independence.

Table 1. CEV-models.

Innovations	α_0	γ	σ	ν	$\log L$
U.S. Treasury Bill					
Normal	0.0009709	0.3755	0.04788		15707.94
t_ν	-0.0002434	0.5377	0.2186	2.01	20688.54
NIBORM3					
Normal	0.0002863	1.61	0.006298		4771.78
t_ν	0.0002419	0.988	0.0921	2.01	8041.42

Note: The constant elasticity volatility (CEV) model of [Longstaff et al., 1992] with normal and student- t innovations.

Table 2. Multifractal models

K	$10^3 \times \alpha_0$	γ	m_0	b	λ_K	σ	$\log L$	BIC	Vuong	HAC-adj.
U.S. Treasury Bill M3										
2	0.2752 (0.247)	0.2824 (0.0106)	1.808 (0.00471)	23.85 (4.92)	0.1391 (0.011)	0.07037 (0.00162)	22309.49	-3.1446	13.573 (< 0.001)	9.091 (< 0.001)
3	0.2052 (0.234)	0.1913 (0.0175)	1.73 (0.0062)	10.62 (1.45)	0.1660 (0.0163)	0.08078 (0.00336)	22676.53	-3.1964	11.618 (< 0.001)	10.138 (< 0.001)
4	0.2276 (0.226)	0.2269 (0.032)	1.661 (0.00674)	9.643 (1.23)	0.2956 (0.0462)	0.07202 (0.00357)	22818.9	-3.2165	8.4076 (< 0.001)	7.562 (< 0.001)
5	0.0643 (0.222)	0.2280 (0.0191)	1.609 (0.00667)	9.804 (1.04)	0.6761 (0.0597)	0.06865 (0.00265)	22901.03	-3.2281	5.036 (< 0.001)	4.805 (< 0.001)
6	0.0375 (0.219)	0.2115 (0.0282)	1.559 (0.00744)	6.577 (0.564)	0.778 (0.0541)	0.06368 (0.00360)	22920.57	-3.2308	4.346 (< 0.001)	4.335 (< 0.001)
7	0.0660 (0.223)	0.1744 (0.0291)	1.527 (0.00752)	5.713 (0.465)	0.7928 (0.0561)	0.08208 (0.00441)	22933.55	-3.2326	3.4883 (< 0.001)	3.112 (0.001)
8	0.0839 (0.215)	0.1575 (0.0260)	1.503 (0.00815)	4.932 (0.355)	0.8474 (0.0509)	0.06956 (0.0036)	22951.89	-3.2352	1.131 (0.129)	0.938 (0.174)
9	0.0693 (0.215)	0.1984 (0.037)	1.462 (0.00724)	3.864 (0.255)	0.931 (0.0396)	0.06029 (0.00443)	22955.80	-3.2358		
NIBORM3										
2	1.156 (0.408)	1.805 (0.04)	1.843 (0.00547)	16.23 (4.68)	0.1016 (0.0116)	0.006092 (0.000516)	8514.45	-2.7250	5.219 (< 0.001)	5.298 (< 0.001)
3	1.041 (0.378)	1.761 (0.0391)	1.772 (0.00685)	18.37 (3.07)	0.4259 (0.0459)	0.005674 (0.000463)	8671.87	-2.7755	3.054 (0.001)	3.093 (< 0.001)
4	1.020 (0.378)	1.482 (0.043)	1.717 (0.00794)	12.43 (1.93)	0.3939 (0.0483)	0.01401 (0.00129)	8760.38	-2.8039	3.424 (< 0.001)	3.588 (< 0.001)
5	1.13 (0.381)	1.933 (0.0635)	1.642 (0.0100)	7.495 (1.04)	0.5455 (0.086)	0.004534 (0.000555)	8779.88	-2.8102	3.292 (< 0.001)	3.265 (0.001)
6	1.079 (0.379)	1.620 (0.0481)	1.590 (0.0107)	5.347 (0.542)	0.6584 (0.0726)	0.00897 (0.000841)	8789.49	-2.8133	2.702 (0.003)	2.866 (0.002)
7	1.006 (0.379)	1.934 (0.0685)	1.537 (0.00926)	4.431 (0.448)	0.9000 (0.0748)	0.003719 (0.000464)	8798.81	-2.8162	1.809 (0.0352)	1.727 (0.042)
8	1.023 (0.380)	1.985 (0.063)	1.500 (0.00933)	3.536 (0.292)	0.9173 (0.0624)	0.003220 (0.000388)	8802.66	-2.8175	1.058 (0.145)	0.940 (0.173)
9	0.998 (0.379)	1.882 (0.112)	1.464 (0.00927)	2.964 (0.216)	0.949 (0.0484)	0.003706 (0.00102)	8804.44	-2.8181		

Note: ML-estimates for the binomial multifractal with estimated standard errors in paranteses. The Vuong-column reports the likelihood ratio statistic with corresponding p -value in brackets [Vuong, 1989]. The null hypothesis is that MSM(9) and MSM(K) have equal Bayesian Information Criteria, with the alternative hypothesis being that the MSM(9) is closer to the true data generating process. The HAC-adjusted version of the Vuong test [Calvet and Fisher, 2004] corrects for heteroschedacity and autocorrelation in the addends.

Table 3. Level-GARCH

$10^3 \times \alpha_0$	γ	$10^5 \times a_0$	a_1	b	ν	$\log L$	BIC
U.S. Treasury Bill M3							
0.0904 (0.225)	0.1699 (0.0496)	0.7092 (0.139)	0.1301 (0.0092)	0.8915 (0.0062)	3.7995 (0.134)	22873.03	-3.2241
NIBORM3							
0.9104 (0.397)	0.9717 (0.0719)	0.4889 (0.158)	0.20993 (0.0209)	0.80838 (0.0135)	3.3305 (0.155)	8783.27	-2.8113

Note: ML estimates for the parameters in the level-GARCH model. Standard deviations are in brackets.

Table 4. Level-EGARCH

$10^3 \times \alpha_0$	γ	a_0	a_1	a_2	b	ν	$\log L$	BIC
U.S. Treasury Bill M3								
0.1039 (0.212)	0.3441 (0.0648)	-0.2360 (0.0165)	-0.03127 (0.00634)	0.2322 (0.0123)	0.9888 (0.00154)	3.866 (0.138)	22936.53	-3.2324
NIBORM3								
1.120 (0.396)	0.9101 (0.0887)	-0.4024 (0.0372)	0.02930 (0.0104)	0.2664 (0.0188)	0.9741 (0.00333)	3.358 (0.156)	8828.90	-2.8245

Note: ML estimates for the parameters in the level-EGARCH model. Standard deviations are in brackets.

Table 5. Jump-diffusion

$10^3 \times \alpha_0$	γ	$10^6 \times a_0$	a_1	b	c	d	τ	$\log L$	BIC
U.S. Treasury Bill M3									
0.464 (0.232)	0.1670 (0.0374)	5.131 (0.867)	0.0931 (0.0055)	0.8773 (0.0062)	-3.751 (0.170)	0.2761 (0.0271)	0.08622 (0.00624)	22597.89	-3.1839
NIBORM3									
1.085 (0.417)	1.872 (0.0892)	0.0436 (0.0198)	0.1291 (0.0099)	0.8301 (0.0111)	-1.883 (0.328)	-0.2119 (0.0417)	0.01127 (0.00216)	8579.61	-2.7431

Note: ML estimates for the parameters in the jump-diffusion model. Standard deviations are in brackets.

Table 6. In-sample model comparison.

Model	dim(θ)	log L	BIC	Vuong	Hac-adj.
U.S. Treasury Bill M3					
MSM(9)	6	22955.80	-3.2358		
GARCH	6	22873.03	-3.2241	-4.064(< 0.001)	-3.477(< 0.001)
EGARCH	7	22936.53	-3.2324	-1.420 (0.078)	-1.361 (0.087)
Jump-diffusion	7	22597.89	-3.1839	-9.878 (< 0.001)	-9.240 (< 0.001)
NIBORM3					
MSM(9)	6	8804.44	-2.8181		
GARCH	6	8783.27	-2.8113	-1.099(0.136)	-1.086(0.139)
EGARCH	7	8828.90	-2.8245	1.231(0.891)	1.470 (0.929)
Jump-diffusion	7	8579.61	-2.7430	-6.964(< 0.001)	-5.765(< 0.001)

Note: The Vuong-column reports the test-statistic for differences in BIC. The null hypothesis is that the multifractal and the alternative model are equally good, with the alternative hypothesis being that the multifractal model performs best. The HAC-adjusted column adjusts for heteroschedacity and autocorrelation in the addends. Corresponding p -values are in brackets.

Table 7. Linear drift

Model	$10^3 \times \alpha_0$	$10^4 \times \alpha_1$
U.S. Treasury Bill M3		
MSM(9)	0.1214 (0.354)	-0.1905 (1.04)
GARCH	0.2821 (0.386)	-0.6731 (1.10)
EGARCH	0.3595 (0.310)	-1.3524 (0.92)
Jump-diffusion	0.7074 (0.397)	-0.8695 (1.15)
NIBORM3		
MSM(9)	2.228 (0.726)	-3.201 (1.60)
GARCH	2.424 (0.746)	-3.986 (1.66)
EGARCH	2.515 (0.430)	-3.626 (0.87)
Jump-diffusion	2.621 (0.788)	-4.053 (1.77)

Note: A linear term in the drift term was added and the parameters estimated using ML. In the TBM3 none of the parameters α_1 are significantly different from zero. Standard deviations are reported in brackets.

References

- T. G. Andersen and J. Lund. Estimating continuous-time stochastic volatility models of the short-term interest rate. *Journal of Econometrics*, 77(2):343 – 377, 1997.
- E. Bacry and J. F. Muzy. Log-Infinately Divisible Multifractal Processes. *Communications in Mathematical Physics*, 236(3):449–475, 2003.
- E. Bacry, J. Delour, and J. F. Muzy. Multifractal random walk. *Physical Review E*, 2001.
- E. Bacry, A. Kozhemyak, and J. F. Muzy. Continuous cascade models for asset returns. *Journal of Economic Dynamics and Control*, 32(1):156–199, 2008.
- T. G. Bali and L. Wu. A comprehensive analysis of the short-term interest-rate dynamics. *Journal of Banking and Finance*, 30(4):1269 – 1290, 2006.
- R. J. Brenner, R. H. Harjes, and K. F. Kroner. Another look at models of the short-term interest rate. *Journal of Financial and Quantitative Analysis*, 31:85–107, 1996.
- L. Calvet and A. Fisher. Forecasting multifractal volatility. *Journal of Econometrics*, 105(1):27 – 58, 2001.
- L. Calvet and A. Fisher. How to Forecast Long-Run Volatility: Regime Switching and the Estimation of Multifractal Processes. *Journal of Financial Econometrics*, 2(1):49–83, 2004.
- D. A. Chapman and N. D. Pearson. Recent advances in estimating term-structure models. *Financial Analysts Journal*, 57:77–95, 2001.
- J. C. Cox, J. E. Ingersoll Jr., and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53:385–407, 1985.
- S. R. Das. The surprise element: jumps in interest rates. *Journal of Econometrics*, 106(1):27 – 65, 2002.
- G. B. Durham. Likelihood-based specification analysis of continuous-time models of the short-term interest rate. *Journal of Financial Economics*, 70(3):463 – 487, 2003.
- Y. Hong, H. Li, and F. Zhao. Out-of-sample performance of discrete-time spot interest rate models. *Journal of Business and Economic Statistics*, 22(4):457–473, 2004.
- M. Johannes. The statistical and economic role of jumps in continuous-time interest rate models. *The Journal of Finance*, 59(1):227–260, 2004.
- K. G. Koedijk, J. A. Nissen, P. C. Schotman, and C. P. Wolff. The Dynamics of Short-Term Interest Rate Volatility Reconsidered. *European Finance Review*, 1(1):105–130, 1997.
- S. Kullback and R. A. Leibler. On information and sufficiency. *Ann. Math. Stat.*, 22:79–86, 1951.
- F. A. Longstaff, K. C. Chan, G. Andrew Karolyi, and A. B. Sanders. An empirical comparison of alternative models of the short-term interest rate. *The Journal of Finance*, 47:1209–1227, 1992.
- O. Løvsletten and M. Rypdal. Approximate Maximum Likelihood Estimation in Multifractal Random Walks. *preprint*, 2011.
- O. Løvsletten. Empirical analysis and stochastic modeling of temporal fluctuations in the norwegian interbank offered rate. Master thesis, written in Norwegian, Autumn 2010. URL <http://complexityandplasmas.net/nordforsk/Papers.html>.
- T. Lux. The markov-switching multifractal model of asset returns : Gmm estimation and linear forecasting of volatility. *Economics Working Papers*, 17, 2006.
- B. Mandelbrot, A. Fisher, and L. Calvet. A Multifractal Model of Asset Returns. *Cowles Foundation for Research in Economics*, 1997.
- Y. Pesin. *Dimension Theory in Dynamical Systems*. Chicago Lectures in Mathematics, 1997.
- M. Rypdal and K. Rypdal. Stochastic modeling of the ae index and its relation to fluctuations in bz of the imf on time scales shorter than substorm duration. *Journal of Geophysical Research*, 115, 2010.
- O. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2): 177 – 188, 1977.
- Q. H. Vuong. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, 57: 307–333, 1989.